WPS3983

Innovation and Export Portfolios*

Bailey Klinger
Center for International Development
Harvard University

Daniel Lederman
Office of the Chief Economist
Latin America and the Caribbean

and

Development Economic Research Group
The World Bank

Abstract: *This paper examines the link between sectoral concentration and overall performance in the search for on-the-frontier innovations, inside-the-frontier innovations, and export booms. We extend the literature by increasing country coverage and the types of search processes considered, and by focusing on the links with overall performance in these search processes. After controlling for the necessary relationships as well as fixed effects at the country/commodity group level, we find a clear negative relationship between the concentration of innovation portfolios and performance: countries that are the most successful in these search processes have their successes spread across a broader range of industries than those with poorer performance. Furthermore, the search for export booms exhibits the least amount of sectoral concentration and path-dependence. These findings suggest that public support for these processes need not be focused in a narrow range of sectors, and modeling of these processes in theoretical work, particularly in the search for export booms, should be of a stochastic flavor.*

Keywords: Diversification, innovation, exports.
JEL Classification: O31, F10

World Bank Policy Research Working Paper 3983, August 2006

We are grateful for insightful discussions with Ricardo Hausmann, J. Humberto López, William F. Maloney, and Marcelo Olarreaga. Financial support from the Regional Studies Program of the Latin America and Caribbean Region of the World Bank is gratefully acknowledged. Send comments to:
dlederman@worldbank.org & bailey_klinger@ksgphd.harvard.edu.

1. Introduction

The purpose of this paper is to investigate the link between sectoral concentration in the search for innovations and export booms and overall performance in these search processes at the country level. In order to inform the modeling of these search processes as well as the debate as to whether national innovation policies should be focused on certain sectors or diversified, we ask the question: Do countries that are high performers in these search processes specialize in a narrow range of sectors, or are their successes spread across a wide range of industries?

Previous research has focused on cross-country patterns and trends in the concentration of patenting activity of OECD countries. We expand this country coverage from the OECD to a much larger worldwide dataset, and extend the focus from only technologically complex on-the-frontier innovation to also include inside-the-frontier innovation (the emergence of new industries, or discoveries) and product-level export booms. All of these forms of innovation are important for growth, involve search across a wide array of potential sectors, and are frequently supported by government in some fashion. As such, this paper assesses the degree to which success in these search processes is focused or diversified in order to inform both public support strategies and modeling in theoretical work.

The question of whether innovation is persistent and concentrated in particular industries or distributed across a wider range of technologies arises from two very different characterizations of innovation. One view, attributed to Schumpeter 1934 and subsequently labeled the Mark I model, is a model that describes innovation as a stochastic process where relatively homogeneous firms fish for innovations in a pool of technological opportunities available to all (Malerba and Orsenigo 1995). According to this view, small entrepreneurs and new firms will succeed in their search for innovations largely through providence, creating temporary monopolies. There will be little persistence at the firm level, and diffuse patterns of innovation at the national level. The consequence of this model of innovation is that, given the stochastic nature of innovation, any public support should be spread as widely as possible across sectors to hedge the government's bets and generate the highest number of draws possible from the technological pool. Examples of models with this random draw feature include Nelson and Winter's (1982) model of innovation as independent technology draws and Eaton et. al.'s (1999) model of patent diffusion.

Another characterization of innovation (attributed to Schumpeter 1942, subsequently labeled the Mark II model) describes it as a cumulative, rather than stochastic, process. In this model of innovation as "creative accumulation", large heterogeneous firms fund R&D laboratories that build new innovations on past advances, using technology that is not commonly available or easily transferable (Malerba and Orsenigo 1995). Under this model, innovation will be characterized by persistent success among a smaller number of established firms. At the national level, this cumulative feature would result in patterns of persistent specialization in certain sectors, and appropriate national policies would target those sectors that will fuel future innovation and growth, rather than wasting resources by spreading them across all sectors. An example of a model with cumulative technological progress is Nelson and Winter's (1982) evolutionary model with the distribution of innovative outcomes determined by past productivity and R&D spending.

There has been significant empirical research examining the degree of support for each model. Malerba and Orsenigo (1995) find robust patterns across countries where some technologies (mechanical technologies and traditional sectors) exhibit Mark I characteristics, while others (chemicals and electronics) exhibit Mark II characteristics. Not surprisingly, those sectors that exhibit hysteresis (Mark II-type sectors) lead to greater specialization at the national level (Malerba, Orsenigo, and Peretto 1997). Other studies at the national level find that although larger countries tend to be less specialized than smaller ones, industrialized countries on the whole exhibit persistent specialization (Pavitt 1988, Cantwell 1989, Eaton et. al. 1999), with specialization increasing over time (Archibugi and Pianta 1992 and 1994). This finding has been disputed by other research that has found no increasing trend of specialization in innovation at the national level, and little persistence across technologies within countries Mancusi (2001, 2003).

These studies have characterized levels and trends of sectoral specialization at the national level given a particular level of innovative performance, but have not examined the relationship between concentration and performance. The superiority of a focused versus diversified innovation strategy would depend on weather or not specialization is associated with better overall performance in innovation.

Another limitation of these studies from the perspective of development is that they have only considered a particular subset of innovation. As defined in Nordhaus (1969), innovation includes the introduction of products and processes that are new to the firm or country, even if not new to the world. Schumpeter’s conception of innovation (1934) also was much broader than that typically considered, as it included the introduction of new goods and the opening of new markets. Unlike the more limited focus of the empirical literature, the theoretical literature considers not only new-to-the-world technologies as innovation, but also the emergence of products that are simply new to a particular country's productive context. This type of innovation, which we will refer to as inside-the-frontier innovation, or discovery, has received increasing attention (Klinger and Lederman 2004, 2006; Hausmann and Rodrik 2003), as for developing countries operating far inside the global technology frontier, it is much more relevant than on-the-frontier innovation.

When considering inside-the-frontier innovation, there is no reason to believe *a priori* that it is the result of cumulative learning or stochastic, especially relative to patentable on-the-frontier innovation. Some models of inside-the-frontier innovation, such as Hausmann and Rodrik 2003, treat productive success in a new product as a purely random draw, which would place it in the Mark I category of innovation. However, one could just as easily argue that unlike new-to-the-world technological development which involves a great deal of creativity and serendipity, inside-the-frontier innovation deals with products and technologies known in other countries, and success is based on knowable comparative advantage and factor endowments. According to this view of the world, inside-the-frontier innovations would be much more concentrated and exhibit the characteristics of Mark II-type innovation. Therefore, in addition to providing a perspective on the category of innovation more relevant to developing countries, the present study will indicate how the nature of innovation changes as one moves inside the technological frontier.

When one looks at export growth at the product level, the aggregate export growth rate masks a reality at the product level where few goods experience rapid accelerations in exports and others lag behind. This is illustrated in Figure 1, which shows a comparison between the deviation from radial growth (the growth rate of the basket as a whole) of individual products in the export basket[1], and the radial growth rate. If all exports grew at the rate of growth of aggregate exports, the value of this deviation would be zero. However, we see that not only is export growth in every country not radial, but those countries that have the fastest export growth overall also have the highest degree of divergence from radial growth at the product level. That is, export growth is characterized by a smaller number of export booms at the product level rather than broad-based radial growth.

Finding those products that will experience booms and fuel future growth is an uncertain search process, similar to the search for globally-novel inventions and the discovery of new products at the national level. Furthermore, this search process is often subsidized by governments through agencies to promote exports abroad, initiatives to facilitate and fund the creation of new clusters, and so on. Do successes in the search for export takeoffs exhibit the characteristics of Mark I innovation, resembling random stochastic draws, or are they cumulative and concentrated in particular successful sectors?

The rest of this paper is organized as follows. Section 2 discusses our estimation strategy designed to determine the extent to which either portfolio diversification or path dependence affect the frequency of on- and inside-the frontier innovations and export

---

[1] Deviation = $\dfrac{\sum_{i=0}^{n} \left| x_{i1} - (1+g)x_{i0} \right|}{X_0}$, where x_{ij} are exports of good i (of which there are n) in period j (0=1995, 1=2002), X_j are total country exports in period j, and g is the compounded annual growth rate of the country's export basket as a whole. This measure increases from 0 as individual product growth rates deviate from the basket's growth rate. This metric comes from work with Ricardo Hausmann.

booms. This section also discusses related literature and data issues. Section 3 presents the econometric results. Section 4 provides results that test the robustness of the previous results through specifications that expand the basic model and control for effective-population scale effects, where effective scale is measured by the number of educated people or non-poor people with the ability to innovate or come up with new commercial ideas. These models thus provide an empirical link between social inequality and poverty and innovation and export booms. Section 5 concludes.

2. Methodology, Related Literature, and Data Issues

Our empirical approach is to estimate the following general econometric model for all three search processes:

(1) $$performance_{c,i,t} = \exp(\alpha + \gamma \times concentration_{c,i,t} + \beta X_{c,t} + \eta_{i,c} + \varepsilon_{c,i,t}),$$

where the subscript c represents countries, i represents industries, and t is a time period. Because the performance measures are integers with a non-negligible frequency of zeros, we use the log-linear formulation estimated by the negative binomial regression. In general, this is the estimation approach used in the vast literature examining counts of patenting activity (e.g., Hausmann et al. 1984; Blundell et al 2000; Blundell et al. 2002; Bosch, Lederman, and Maloney 2005), as well as in emerging empirical literature on the determinants of the frequency of export discoveries (Klinger and Lederman 2004, 2006).[2]

[2] There is Monte Carlo evidence showing that even if the frequency of zeroes is negligible, count-data methods are preferable to linear estimators as long as the data generation process produces heteroskedasticity such that the variance of the errors is correlated with the conditional mean – see Santos Silva and Tereyro (forthcoming). The Negative Binomial estimator allows for over dispersion of the errors with respect to the conditional mean. Specification tests for our applications suggested that there is over dispersion, and thus the Negative Binomial is preferred over the Poisson estimator.

In all of our models we control for time-invariant unobserved heterogeneity across country-industries, which is represented in equation (1) by the parameter $\eta_{c,i}$. To control for this source of hetergoneity across countries and industries, we employ Blundell et al.'s (2002) Pre-Sample Mean (PSM) estimator, whereby historical innovation counts are used to control for time-invariant characteristics that can affect the frequency of innovation. In the case of our models explaining the variance in patenting activity across country-industries, where we utilize more than one time period, we also control for time effects and for dynamic effects through the inclusion of the lagged dependent variable. Unfortunately, as discussed by Blundell et al. (2002), in the PSM estimator, the historical patent counts used to control for time-invariant characteristics tends to absorb some of the persistence that would be captured by the lagged dependent variable. This is an important feature of these models for inferring the extent to which there is cumulative learning in the search process, because the historical performance measures can be thus interpreted as a measure of persistence, as well as of the presence of fixed effects, which would nevertheless produce persistence in the performance measures. Consequently for inference, we interpret the magnitudes and significance of the PSM's historical performance variables as evidence of cumulative learning.

Our measure of concentration is the Gini coefficient of the Technological Revealed Comparative Advantage (TRCA) index. This measure of concentration has been used in research examining patterns of specialization in patenting (Mancusi 2001, Amiti 1999). While the Gini coefficient traditionally compares a realized distribution to a theoretical uniform distribution, the Gini coefficient of the TRCA compares a country's distribution (in our case across industries and commodity groups) to the global distribution across those sectors. This modification is necessary because unlike the Gini coefficient for income, for example, where the benchmark is a completely uniform distribution of wealth, there is no reason to assume that the global distribution of patents is uniform across industries.

The concentration measure is calculated as follows. First the TRCA index constructed using counts of patents/discoveries/booms *C* for country *i* in industry *j*:

$$(2) \qquad TRCA_{ij} = \frac{C_{ij} / \Sigma_j C_{ij}}{\Sigma_i C_{ij} / \Sigma_{ij} Cij}$$

The numerator is the ratio of country counts in industry *j* to country counts in all industries. The denominator is the equivalent at the global level: worldwide counts in industry *j* to worldwide counts in all industries. As in Mancusi (2001), the Lorenz curves plot the cumulative numerator on the vertical axis against the cumulative denominator on the horizontal axis after ordering observations in ascending order by the value of the TRCA index. The Gini index is, as usual, twice the area between the 45-degree line and the Lorenz curve. If a particular country had counts of discoveries, patents, or booms spread across sectors in the same pattern as the worldwide distribution, the value of this index would be zero. However, as a country deviates from the global norm and concentrates in particular sectors, this value approaches 1. Therefore, finding a negative value for the coefficient on concentration in equation (1), γ, indicates that greater success occurs when specialization decreases, whereas a positive estimated value of γ suggests that success is associated with higher specialization.

A common control variable for the three types of search process investigated herein is the size of a country's population, which is used to capture scale effects affecting the frequency of innovations across countries. That is, countries with larger populations are expected to have higher counts of innovations than smaller countries. The following paragraphs provide further details about the data sources and control variables for each of the three search processes.

On-the-Frontier Innovation

Our indicator of on-the-frontier innovation performance is patent grants by the U.S. Patent and Trademark Office (USPTO) collected by Lederman and Saenz (2005). The data are available from 1963 to 2001; however plant and design patents are only available after 1978. The period 1963 to 1981 is used to identify historical patent counts, and the 20-year period 1982-2001 is divided into four 5-year panels. The sum of patents within each panel is the indicator of performance in on-the-frontier innovation for that period. The Gini index is calculated as described above across the 43 industry categories, plus design and plant categories, available from the USPTO.

Because the incentive to patent an invention in the United States will largely depend on the importance of that export market to the particular country, we control for per-capita exports to the United States during the period. In addition, we control for investment in research and development (R&D) per capita which has been shown to be a highly significant determinant of patenting activity by firms as well as across countries (Hausmann et al. 1984; Lederman and Maloney 2003; Blundell et al 2000; Bosch et al. 2005, among others), and population to capture the effects of scale. Distance to the global technological frontier is captured by controlling for GDP per capita in quadratic form, and period indicators are added to capture temporal trends. Finally, we control for the historical level of patenting activity to capture country fixed-effects in patenting, and the previous period's patenting activity to consider the path dependence that would be present if there were hysteresis in patenting. Using this lagged variable provides us with three periods per country in the sample.

Inside-the-Frontier Innovation

Performance in inside-the-frontier innovation is measured by discovery counts as described in Klinger and Lederman (2006), available for a cross-section of countries during the period 1997-2002. The unit of observation is the Leamer (1984) commodity group in each country, and therefore the performance measure is discovery counts in the country/commodity cluster. Discovery counts are first aggregated to the SITC 3-digit level, and the Gini coefficient for each country/Leamer commodity cluster was calculated as described above across the SITC 3-digit clusters comprising the commodity group.

The controls comprising the X vector in this case follow from Klinger and Lederman (2006). First, as above, we control for the historical level of discovery activity in the country/commodity group to control for country/commodity group fixed effects. Unfortunately, discovery data is only available in a cross-section, rather than in multiple panels as is the case with our patent data. As such, we are unable to include a lagged variable that allows us to distinguish between fixed country effects and persistence that from path-dependent innovation. In the case of discoveries and export booms, both of these effects will be captured in the historical variable.

It is also necessary to control for GDP per capita (in quadratic form) to capture the discovery curve identified in Klinger and Lederman (2006) and capture distance to the global technological frontier. Finally, we include commodity cluster dummy variables to capture the differences between the 10 Leamer commodity clusters, and country population to account for scale.

Export Booms

Export booms are identified using the identical product-level export data used to identify discoveries. A product export boom is defined as a product experiencing at least 10% growth in at least 6 years between 1997 and 2003. There are 3782 product-level export booms identified in the data, listed by country in the Appendix. As with the discovery data, counts of these booms are aggregated to the SITC 3-digit level, and the Gini coefficient for the Leamer commodity cluster is calculated, providing our measure of concentration.

The control variables included in our examination of export booms are historical boom counts for the country/Leamer commodity group to capture country/commodity group fixed effects and persistence, GDP per capita (in quadratic form) to capture the possibility of more export accelerations in poorer countries due to convergence, population level to control for scale, and commodity cluster indicator variables.

3. Results

Figures 2, 3, and 4 show the relationship of performance, measured by total counts, to concentration[3]. Among all three search processes (on-the-frontier innovation, discoveries, and export booms), there is a clear negative correlation between concentration and performance.

[3] For illustrative purposes, performance and concentration at the country level across commodity groups is shown. The estimation, however, is at the country/commodity group level, as described in the methodology section.

Table 1 shows the results of estimating equation (1) for each of the three search processes: on-the-frontier innovation, inside-the-frontier innovation, and export booms.

The coefficients on income are as expected. The discovery curve identified in Klinger and Lederman (2006) is maintained, as is the exponential relationship between level of development and on-the-frontier innovation found in Lederman and Maloney (2003). The form of the quadratic relationship between GDP per capita and export booms is similar to the discovery curve, indicating that poorer countries have a higher frequency of such booms, which is consistent with the convergence hypothesis.

Scale effects captured by the number of people in the country that are potentially searching for new innovations and products is positive and significant for all three search processes. This is consistent with intuition: *ceterius paribus*, a larger country will have more potential innovators and entrepreneurs engaged in the search processes under consideration, and therefore are expected to have a higher number of successes.[4]

Our control for fixed effects, historical performance, is found to be positive and significant for both on-the-frontier and inside-the-frontier innovation, but not in the search for export booms. This suggests greater hysteresis, or path dependence, in innovation compared to export booms, or alternatively that time-invariant country and sector characteristics are not significant determinants of the frequency of export booms. Finally, for on-the-frontier innovation, expenditure on R&D is found to be highly significant and positive, as are exports per capita to the United States, suggesting we are

[4] Admittedly, scale effects could be captured by the log of GDP. In fact, when this variable is used instead of population, the coefficients of GDP do capture the effect of scale. But both cannot be included simultaneously with the log of GDP per capita in the same regression. Our preferred specification is with log of population, because this approach naturally leads in the estimations concerning the role of poverty rates and access to education by the population, which are discussed in section 4 below.

properly controlling for the fact that on-the-frontier innovation data is limited to patents within the United States.

One surprising result shown in Table 1 is that export growth net of non-export GDP growth, which is meant to capture the relative profitability in searching for new exports, is insignificant. In Klinger and Lederman (2006), this variable was found to be positively related to discovery, particularly as barriers to entry increase, suggesting that fears of imitation dull the incentives to experiment and retard discovery activity. In order to verify the robustness of this finding, we split the sample into low barrier and high barrier countries and find that returns do in fact have a positive relationship with discovery in high-barriers countries. Furthermore, the relationship between net export growth and discovery was the only relationship to be affected by differences in barriers to entry. The other variables shown in Table 1 did not significantly change when the sample was split according to the Klinger and Lederman (2006) barriers index, confirming the conclusions therein.

In addition, we extended the analysis of barriers to on-the-frontier innovation by examining the effects of the intellectual property rights (IPR) regime, taken from Park (2001), on the relationships of interest. As expected, the quality of the IPR regime enters the full estimation of patent performance as positive and significant. Furthermore, when the sample is split into the best and worst halves with respect to quality of the IPR regime, the elasiticty of R&D spending on patenting changes significantly, becoming much higher in countries with better IPR regimes. This conforms with intuition: just as discovery is higher in countries where innovators are able to appropriate the returns of their discovery, the value of the search for on-the-frontier innovations (measured by

R&D spending) is higher in environments where the returns from such innovations are protected in the domestic market.

After controlling for the necessary variables, we see a consistent negative relationship between concentration and overall innovative performance. When controlling for fixed effects via the historical performance variable, countries (in the case of patents) and country's commodity clusters (in the case of discoveries and booms) perform better in these search processes when patents, discoveries, and export booms are spread across a wider array of products within the broad Leamer sectors rather than concentrated in a few particular products within these broad industries. Performance in the search for export booms seems particularly closely linked to a diversified search, strengthening the finding with respect to hysteresis from the historical variable.

4. Robustness

Table 1 shows that scale is positively related to success in these three search processes, consistent with the idea that more potential innovators and entrepreneurs will result in higher observed cases of innovation and export booms. However, what is important should not be the absolute size of the population, but the size of the relevant population, namely those who are able to innovate by adapting existing products to the local productive environment or invent new products. All else equal, a country featuring high levels of inequality in income or education will have a smaller relevant population.

One way to examine this relationship is to add to the vector of explanatory variables a measure of this relevance and to interact it with population levels. Table 2 shows the results for both types of innovation when we add either the poverty rate or the percentage of the labor force with secondary education and interact it with population. In all four cases, we find that the data confirms the idea that it is the size of the relevant population that matters for innovation, and highly unequal distributions of income resulting in high poverty rates or low education coverage result in lower performance in these search processes. This finding does not, however, extend to the search for export booms, for which measures of poverty and their interaction with population are not significant.

5. Conclusions

According to our measure of concentration, the evidence clearly indicates that countries with the best performance in the search for on-the-frontier innovations, inside-the-frontier innovations, and export booms are also those with the lowest degree of specialization at the sector/industry level in such processes. This is particularly true of export booms, which also exhibit the least amount of hysteresis, suggesting that they are much less cumulative than the innovation processes. Comparing all three search processes to one another, on-the-frontier innovation proxied by patents seems to exhibit the weakest link between diversification and performance, but also a lower level of hysteresis than inside-the-frontier innovation.

These findings suggest that it may not be advisable for countries to “put all their eggs in one basket” when it comes to supporting these three search processes. Particularly in the search for export booms, those sectors with many successes yesterday need not have many successes today. Furthermore, sectoral concentration is closely linked with fewer booms and innovations at the national level. When modeling such processes, a stochastic rather than cumulative flavor would be more in line with the experience of countries during the 1990s. Last but not least, national poverty and access to secondary education might be important determinants of innovation, both on- and inside- the global technological frontier. These relationships could thus also explain why poverty by itself might retard economic growth, namely by thwarting innovation.

Although our main empirical findings seem to be quite robust, no econometric approach is without faults. In our case, much research remains ahead to ascertain whether the diversification of innovation across industries *causes* greater overall innovation, much in the same way as diversified financial portfolios tend to generate higher returns. The evidence presented above suggests that this is a worthwhile research agenda, particularly because the partial correlations between innovation diversification and over innovation tend to be large and statistically significant.

6. References

Amiti, M. 1999. Specialization Patterns in Europe." *Review of World Economics* 135(4): 573-93.

Archibugi, D., and M. Pianta. 1992. "Specialization and Size of Technological Activities in Industrial Countries: The Analysis of Patent Data." *Research Policy* 21: 79-93.

Archibugi, D., and M. Pianta. 1994. "Aggregate convergence and sectoral specialization in innovation." *Evolutionary Economics* 4(1): 17-34.

Blundell, Richard, Rachel Griffith, and John Van Reenen. 1995. "Dynamic Count Data Models of Technological Innovation." *Economic Journal* 105: 333-344.

Blundell, Richard, Rachel Griffith, and Frank Windmeijer. 2002. "Individual Effects and Dynamics in Count Data Models." *Journal of Econometrics* 108: 113-131.

Bosch, M., D. Lederman, and W.F. Maloney. 2005. "Patenting and Research and Development: A Global View." Policy Research Working Paper 3739, The World Bank, Washington, DC.

Breschi, S., F. Malerba and L. Orsenigo. 2000. "Technological Regimes and Schumpeterian Patterns of Innovation." *The Economic Journal* 110: 388-410.

Cantwell, J. 1989. *Technological Innovation and Multinational Corporations*. Basil Blackwell, Cambridge MA.

Deininger, K., and L. Squire. 1996. "Measuring Income Inequality: a New Database." Mimeo. World Bank Development Economics Research Group. Washington DC.

Eaton, J. R. Evenson, S. Kortum, P. Marino, and J. Putnam. 1999. "Technological specialization in international patenting." In *Development, duality, and the international economic regime: essays in honor of Gustav Ranis* (G. Saxonhouse & T. Srinivasan, eds.). University of Michigan Press, Ann Arbor MI.

Hausman, Jerry, Bronwyn H. Hall, and Zvi Griliches. 1984. "Econometric Models for Count Data with an Application to the Patents-R&D Relationship." *Econometrica* 52(4): 909-938.

Hausmann, R., and D. Rodrik. 2003. "Economic development as self-discovery." *Journal of Development Economics* 72: 603-633.

Klinger, B., and D. Lederman. 2004. "Development and discovery: an empirical exploration of 'new' products. Policy Research Working Paper 3450, The World Bank, Washington DC.

Klinger, B., and D. Lederman. 2006. “Diversification, innovation and imitation inside the global technological frontier.” Policy Research Working Paper 3872, World Bank, Washington, DC.

Lederman, D., and W. Maloney. 2003. “R&D and Development.” World Bank Policy Research Working Paper #3024. Washington DC.

Lederman, D., and L. Saenz. 2005. “Innovation around the World: A Cross-Country Database of Innovation Indicators.” Mimeo. Office of the Chief Economist for Latin America and the Caribbean, World Bank. Washington, DC.

Leamer, E.E. 1984. *Sources of Comparative Advantage: Theory and Evidence*. Cambridge MA: MIT Press.

López, J.H., and L. Servén. 2006. “A Normal Relationship? Poverty, Growth, and Inequality.” Policy Research Working Paper 3814, The World Bank, Washington, DC.

Malerba, F., and L. Orsenigo. 1995. “Schumpeterian Patterns of Innovation Are Technology-Specific.” *Research Policy* 25: 451-478.

Malerba, F., L. Orsenigo, and P. Peretto. 1997. “Persistence of Innovative Activities, Sectoral Patterns of Innovation and International Technological Specialization.” *International Journal of Industrial Organization* 15:801-826.

Mancusi, M. 2001. “Technological Specialization in Industrial Countries: Patterns and Dynamics.” *Review of World Economics* 137(4): 593-621.

Mancusi, M. 2003. “Geographical Concentration and the Dynamics of Countries’ Specialization in Technologies.” *Economics of Innovation and New Technology* 12(3): 269-291.

Nelson, R., and S. Winter. 1982. *An Evolutionary Theory of Economic Change*. Harvard University Press, Cambridge MA.

Nordhaus, William D. 1969. *Invention, Growth and Welfare*. Cambridge, Mass.: MIT Press.

Park, W. 2001. “Intellectual Property and Patent Regimes.” Chapter 4 in *Economic Freedom of the World: 2001 Annual Report*. Freedom House: New York.

Pavitt, K. 1988. “International patterns of technological accumulation.” In *Strategies in Global Competition*, N Hood & J. Vahlne (eds.). Croom Helm, New York NY.

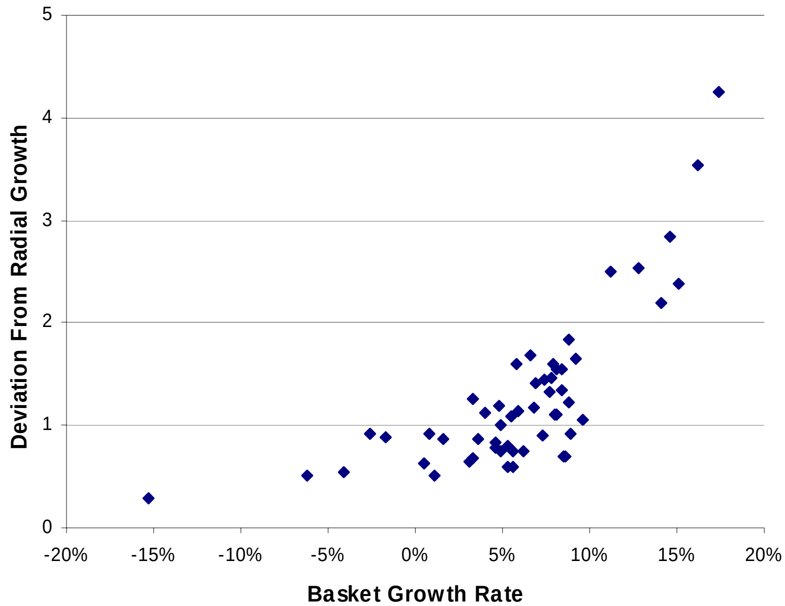
Santos Silva, J.M.C., and S. Tereyro. Forthcoming. “The Log of Gravity.” *Review of Economics and Statistics.*

Schumpeter, J. 1942 *Capitalism, Socialism and Democracy*. Harper, New York NY.

Schumpeter, J. 1934. *The Theory of Economic Development*. Harvard Economic Studies, Cambridge MA.

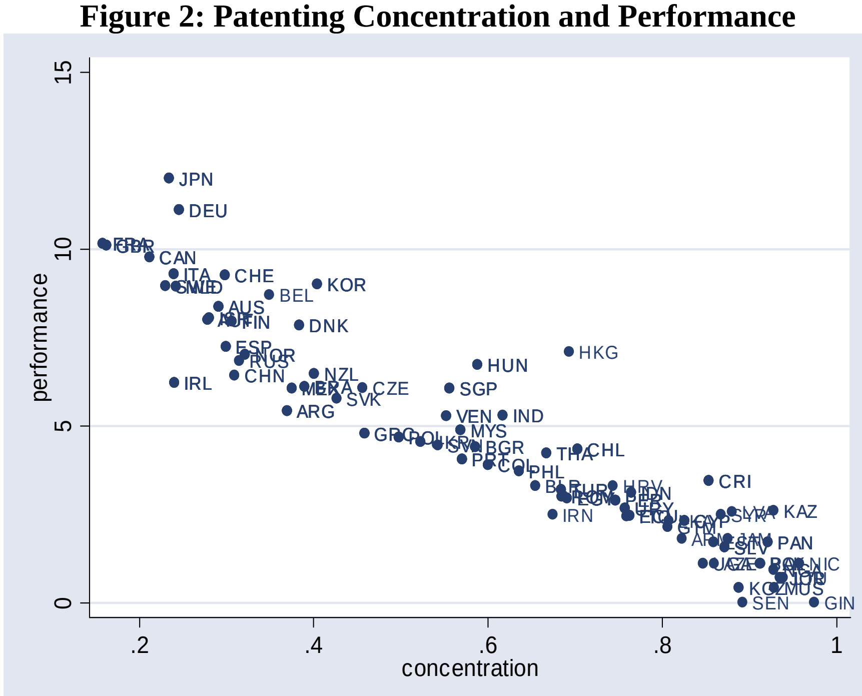
Figures

Figure 1: Export Growth and Deviations from Radial Growth Rates



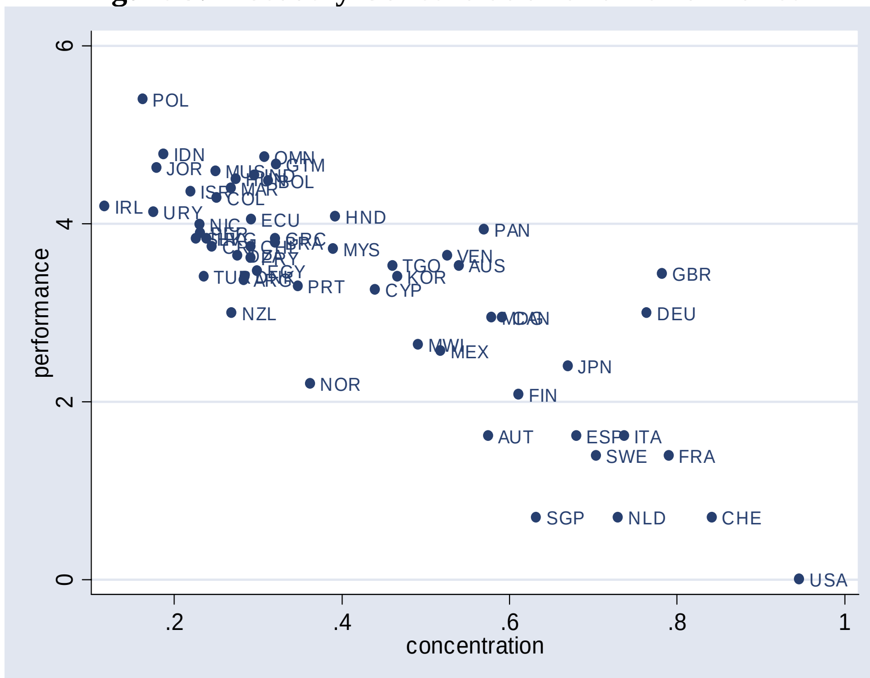
Source: Authors' calculations based on COMTRADE

Figure 2: Patenting Concentration and Performance



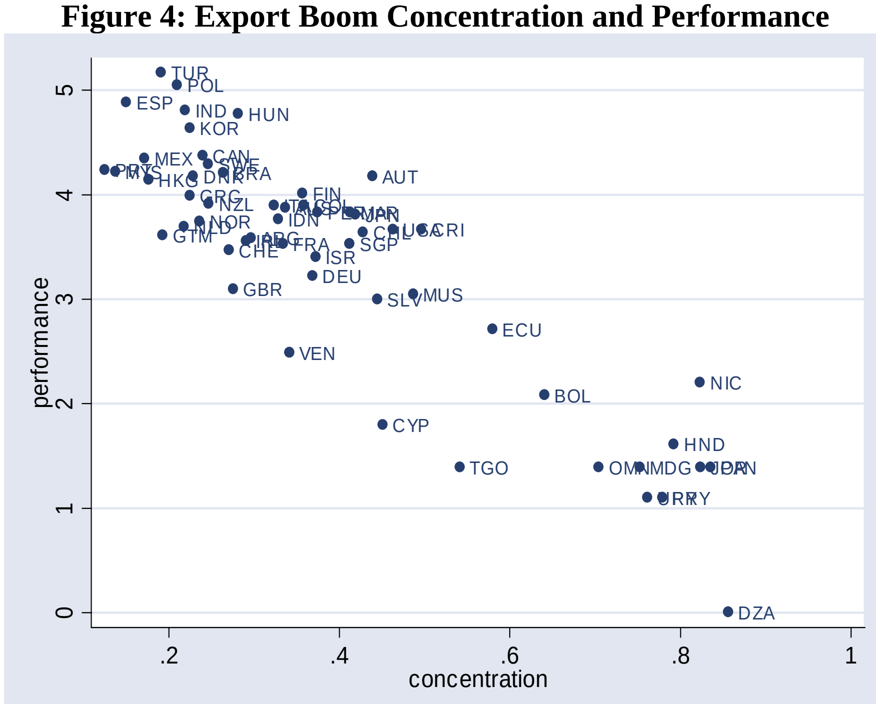
Source: Authors' calculations. Performance is log of average counts across observations (given the implicit log-linear specification of the negative binomial regressions presented below), concentration is the average Gini across observations, only observations included in estimation are shown.

Figure 3: Discovery Concentration and Performance



Source: Authors' calculations. Performance is log of discovery counts (given the implicit log-linear specification of the negative binomial regressions presented below), concentration is the Gini across Leamer commodity groups, only observations included in estimation are shown.

Figure 4: Export Boom Concentration and Performance



Source: Authors' calculations. Performance is log of boom counts (given the implicit log-linear specification of the negative binomial regressions presented below), concentration is the Gini across Leamer commodity groups, only observations included in estimation are shown.

Tables

Table 1: Estimation Results: Innovation Portfolios and Path Dependence as Determinants of Innovation

	Discoveries	**Patents**	**Export Booms**
Gini of innovations	-2.874	-2.522	-3.052
	(14.96)***	(4.76)***	(15.89)***
ln(historical performance)	0.175	0.064	-0.070
	(2.95)***	(2.05)**	(1.52)
ln(GDP per capita)	2.239	-2.393	3.028
	(3.55)***	(1.46)	(3.27)***
ln(GDP per capita)^2	-0.135	0.194	-0.163
	(3.73)***	(2.01)**	(3.20)***
ln(population)	0.038	0.696	0.139
	(1.70)*	(11.06)***	(5.81)***
Net export growth	-0.009		
	(1.44)		
ln(R&D per capita)		0.437	
		(7.16)***	
ln(exports to the US per capita)		0.159	
		(3.22)***	
ln(previous period patents)		-0.005	
		(0.22)	
Constant	-6.947	-1.754	-13.834
	(2.50)**	(0.24)	(3.19)***
Observations	391	179	385

Absolute value of z statistics in parentheses. Industry dummies (for booms & discoveries) and period dummies (for patents) omitted for brevity. Full estimation results can be found in the Appendix.
* significant at 10%; ** at 5%; * at 1%

Table 2: Scale and Inequality

	(1)	(2)	(3)	(4)
	Discoveries	Discoveries	Patents	Patents
ln(GDP per capita)	3.653	4.161	-0.523	-0.165
	(3.38)***	(2.11)**	(0.24)	(0.07)
ln(GDP per capita)^2	-0.227	-0.251	0.056	0.030
	(3.39)***	(2.29)**	(0.45)	(0.20)
Gini of innovations	-3.802	-2.767	-1.539	-2.039
	(8.56)***	(4.75)***	(1.98)**	(3.15)***
ln(historical performance)			0.135	0.045
			(2.51)**	(1.22)
ln(Population) (a)	-0.074	-0248	0.601	0.629
	(1.17)	(1.35)	(5.73)***	(4.99)***
ln(Poverty) (b)	0.378		0.315	
	(2.30)**		(2.07)	
(a) x (b)	-0.025		-0.023	
	(2.36)**		(2.36)**	
Labor force with secondary ed'n (c)		-0143		-0.072
		(2.35)**		(1.55)
(a) x (c)		0009		0.005
		(2.36)**		(1.83)*
ln(R&D per capita)			0.423	0.712
			(3.33)***	(7.16)***
ln(exports to the US per capita)			0.282	0.231
			(4.24)***	(3.64)***
ln(previous period patents)			0.005	-0.024
			(0.16)	(1.07)
Constant	-8.257	-8.432	-7.758	-8.996
	(1.90)*	(0.85)	(0.75)	(0.79)
Observations	70	44	132	115

Absolute value of z statistics in parentheses. Period dummies were also included for patent estimations (columns 3 and 4). Ln(historical performance) was not found to be significant for either (1) or (2) and did not affect the other coefficients so it was excluded to increase sample size.
* significant at 10%; ** at 5%; *** at 1%.
Notes:

6. Appendix

Export Booms By Country

Argentina	36	United Kingdom	22	Nicaragua	9
Australia	48	Greece	54	Netherlands	40
Austria	65	Guatemala	37	Norway	42
Bolivia	8	Hong Kong, China	63	New Zealand	50
Brazil	67	Honduras	5	Oman	4
Central African Republic	0	Croatia	66	Panama	4
Canada	79	Hungary	118	Peru	46
Switzerland	32	Indonesia	43	Poland	155
Chile	38	India	122	Portugal	69
China	473	Ireland	35	Paraguay	3
Cote d'Ivoire	0	Israel	30	Romania	156
Colombia	49	Italy	49	Sudan	1
Costa Rica	39	Jordan	4	Singapore	34
Cyprus	6	Japan	45	El Salvador	20
Czech Republic	169	Korea, Rep.	103	Slovak Republic	113
Germany	25	Latvia	72	Slovenia	68
Denmark	65	Morocco	46	Sweden	73
Algeria	1	Moldova	0	Togo	4
Ecuador	15	Madagascar	4	Turkey	175
Egypt, Arab Rep.	0	Mexico	77	Uganda	0
Spain	132	Macedonia, FYR	17	Uruguay	3
Estonia	121	Mauritius	21	United States	39
Finland	55	Malawi	0	Venezuela	12
France	34	Malaysia	68		
Gabon	4	Niger	0		

Complete Estimation Results

Patents:

	Patents
Gini	-2.522
	(4.76)***
lnPopulation	0.696
	(11.06)***
ln(GDP per capita)	-2.393
	(1.46)
ln(GDP per capita)^2	0.194
	(2.01)**
ln(R&D per capita)	0.437
	(7.16)***
ln(exports to the US per capita)	0.159
	(3.22)***
ln(previous period patents)	-0.005
	(0.22)
ln(historical patents)	0.064
	(2.05)**

Period 3 dummy	0.075
	(0.49)
Period 4 dummy	0.379
	(2.33)**
Constant	-1.754
	(0.24)
Observations	179

Absolute value of z statistics in parentheses
* significant at 10%; ** at 5%; *** at 1%

Discoveries:

	Discoveries
Gini	-2.874
	(14.96)***
Net export growth	-0.009
	(1.44)
ln(historical discoveries)	0.175
	(2.95)***
ln(GDP per capita)	2.239
	(3.55)***
ln(GDP per capita)^2	-0.135
	(3.73)***
ln(population)	0.038
	(1.70)*
commodity group 3	-0.478
	(3.00)***
commodity group 4	-0.461
	(2.85)***
commodity group 5	0.277
	(1.91)*
commodity group 6	0.068
	(0.45)
commodity group 7	0.501
	(3.71)***
commodity group 8	0.458
	(3.44)***
commodity group 9	0.327
	(2.43)**
commodity group 10	0.735
	(5.82)***
Constant	-6.947
	(2.50)**
Observations	391

Absolute value of z statistics in parentheses
* significant at 10%; ** at 5%; * at 1%

Export Booms:

	Export Booms
Gini	-3.052

	(15.89)***
ln(historical booms)	-0.070
	(1.52)
ln(GDP per capita)	3.028
	(3.27)***
ln(GDP per capita)^2	-0.163
	(3.20)***
ln(population)	0.139
	(5.81)***
commodity group 2	0.182
	(1.06)
commodity group 4	0.213
	(1.42)
commodity group 5	0.497
	(3.10)***
commodity group 6	0.726
	(4.10)***
commodity group 7	1.338
	(9.59)***
commodity group 8	1.025
	(7.23)***
commodity group 9	0.800
	(5.65)***
commodity group 10	1.333
	(9.40)***
Constant	-13.834
	(3.19)***
Observations	385

Absolute value of z statistics in parentheses
* significant at 10%; ** at 5%; * at 1%

Data Definitions and Sources

Variable Name	Description	Units	Year(s) Used	Transformation	Source
Patents	Patents granted by the USPTO	Sum of Counts in the period	1963-1981 (historical) 1982-1986 (period 2) 1987-1991 (period 3) 1992-1996 (period 4) 1997-2001 (period 5)	None	Lederman and Saenz (2003)
Discoveries	Counts of discoveries in the country/commodity group	Counts	1997-2002	None	Klinger and Lederman (2005)
Export Booms	Counts of export booms (products with growth greater than 10% in at least 6	Counts	1997-2003	None	COMTRADE

	years between 1997 & 2003.				
US exports	Exports to the United States	Average exports to the United States during the period	1982-1986 (period 2) 1987-1991 (period 3) 1992-1996 (period 4) 1997-2001 (period 5)	Adjusted for inflation using US PPI (all commodities). log.	US PPI: St. Louis FRED. Export Data: IMF Direction of Trade Statistics
R&D expenditure	Expenditures on R&D	Average expenditure during the period (constant 1995 USD)	1982-1986 (period 2) 1987-1991 (period 3) 1992-1996 (period 4) 1997-2001 (period 5)	Log**	Lederman and Saenz (2003)
ln(GDP Per Capita)	Natural log of real GDP per capita (PPP)	2000 PPP Constant Prices	1995 (Table 1 uses all years)	log	World Bank WDI
Population	Population	Count	1995	log	World Bank WDI
Net Export Growth	Growth rate of exports of that Leamer commodity group less average of annual growth rates of non-export GDP 1994 – 2003	decimal form	1994-2003	None	COMTRADE & World Bank WDI
ln(historical booms)	Historical Boom counts***	Total counts between 1984 & 1993	1984-1993	log**	COMTRADE
ln(historical discoveries)	Historical Discovery counts*	Total counts between 1984 & 1993	1984-1993	log**	COMTRADE
Factor Endowments	Average value of net exports per capita between 1989 & 1993 per capita	Current dollars per capita	1989-1993	Net exports for each year divided by that year's population.	COMTRADE & World Bank WDI
Income Gini	Average value of Gini	Index	1985-1995	None	Deininger & Squire (1996) (high-quality national Ginis only)
Poverty Count	Average percentage of the population living on less than $1/day. Calculated using the GDP per capita and income	Percentage	1985-1995	None	World Bank WDI

	Gini indexes under the assumption of log normality – see López and Servén (2006).				
Labor force with secondary ed'n	Average percentage of the total labor force with at least secondary education	Percentage	1990-1995	None	World Bank WDI

*Historical discovery counts are identified using export data from 1970 onwards at the SITCr1 3-digit level. The filter identifies a discovery in the year it first appears as an export greater than 0. The period 1974-1983 is used to create baseline of existing exports, the period 1984-1993 to generate counts of discoveries. The filter drops countries from the sample missing more than 7 years of data in the 1974-1983 period (to ensure at least three years of data exist to identify existing exports) and more than 5 years of data in the 1984-1993 period.

**Before taking logs, 1 added to each to keep observations of 0 in the sample

***Historical boom counts are identified using export data from 1970 onwards at the SITCr1 3-digit level. The filter identifies a discovery in the year it first appears as an export greater than 0. A boom is identified when there are six or more years of growth greater than 10% during the period 1984-1993.